



# Stage migration resulting from inadequate number of examined lymph nodes impacts prognosis in stage II colon cancer after radical surgery

Di Xie<sup>1</sup> · Xiangping Song<sup>2</sup> · Lingling Tong<sup>3</sup>

Accepted: 29 October 2020 / Published online: 10 November 2020  
© Springer-Verlag GmbH Germany, part of Springer Nature 2020

## Abstract

**Purpose** We evaluated the impact of examined lymph node (ELN) number on the prognosis of stage II colon cancer after radical surgery and developed a novel prognostic scoring system by combining primary tumor extension (pT) and ELN number for reclassification of stage II colon cancer.

**Methods** Three cohorts of patients diagnosed with colon cancer between 2004 and 2010 were identified from the Surveillance, Epidemiology, and End Results (SEER) database. Univariate and multivariate analyses were performed to evaluate the relationship between factors and patients' survival including cause-specific survival (CSS) and overall survival (OS). Survival curves from subgroups were plotted by the Kaplan-Meier method and compared by the log-rank test.

**Results** Cohort 1 and cohort 2 consisted of 13,960 and 5312 stage II colon cancer patients, respectively. Cohort 3 consisted of 4713 stage III patients. Factors including ELN, age, and pT were found to be associated with patients' survival in cohorts 1 and 2. Patients who were older or with smaller tumors were more likely to experience inadequate ELN. Patients with a higher score, as calculated by the novel scoring system, showed worse survival. Compared with stage III colon cancer patients, stage II patients with high scores had a comparable or even worse survival than stage IIIA and IIIB patients.

**Conclusion** Inadequate ELN leads to understaging in stage II colon cancer and predicts inferior prognosis. Our analyses show that the novel prognostic scoring system, consisting of combined pT and ELN, quantified stage migration effect and can be applied to the reclassification of stage II colon cancer.

**Keywords** Colon cancer · Examined lymph nodes · Prognostic score system · SEER

## Introduction

Colon cancer is the fourth most commonly diagnosed and fifth cause of cancer-related deaths worldwide [1]. In the USA, approximately 38% of colon cancer patients diagnosed between 2006 and 2012 had localized (stage I and II) colon

cancer and a 5-year survival rate of about 90% better than regional (stage III) colon cancer patients [2]. The standard treatment for non-metastatic colon cancer is radical surgery with or without adjuvant chemotherapy. Adjuvant chemotherapy is recommended for high-risk patients with stage II colon cancer [3, 4]. Great efforts have been made to identify high-risk patients, but there is still no standardized method. Adjuvant chemotherapy can improve the survival of patients with stage III colon cancer but whether it also benefits stage II colon cancer patients remains a contentious matter [5].

TNM staging classification plays an important role in the stratification of patients' prognosis, which guides treatment strategy. Though the prognosis of stage II colon cancer patients is generally better than stage III colon cancer patients, there is still a high level of variability [6–8]. Recently, several studies have reported that a higher number of examined lymph node (ELN) predicted better prognosis among different cancers [9–11]. ELN has also been linked to the prognosis of

---

✉ Xiangping Song  
song-898785@hotmail.com

<sup>1</sup> Department of Neurology, The Third Xiangya Hospital, Central South University, Changsha, People's Republic of China

<sup>2</sup> Department of Gastrointestinal Surgery, Xiangya Hospital, Central South University, Changsha, People's Republic of China

<sup>3</sup> Department of Epidemiology and Health Statistics, Xiangya School of Public Health, Central South University, Changsha, People's Republic of China

colorectal cancer patients [12, 13]. Precise staging is key to execute appropriate adjuvant therapy for colorectal cancer, where the number of positive regional lymph nodes is a critical factor. However, an insufficient number of ELN may lead to inaccurate assessment for lymph node involvement [9, 10]. To identify the potential stage III colon cancer from stage II cancer is of great clinical significance.

According to the TNM staging method, only tumor invasion depth (pT), and not regional lymph nodes, is taken into consideration for prognosis stratification of stage II colon cancer. Although the relationship between ELN levels and colorectal cancer prognosis has been reported, the impact of ELN on stage migration in stage II colon cancer has rarely been studied [12–16]. To address this issue, in our present study we combined the ELN number and pT to generate a novel scoring system for reclassification of stage II colon cancer.

## Methods

### Data source

With granted access to the Surveillance, Epidemiology, and End Results (SEER) Program database (<https://seer.cancer.gov/>), the SEER 1975–2016 Research Plus Additional Custom Treatment Data released in April 2019 was downloaded for analysis. There are two files titled “YR1975\_2016.SEER9” (dataset 1) and “YR1992\_2016.SJ\_LX\_RG\_AK” (dataset 2), respectively. Dataset 1 contains the SEER November 2018 Research Data files from nine SEER registries for 1975–2016 and dataset 2 from another four registries.

### Patients included and cohort selection

The raw data was processed and saved by Excel Software. There were 147 variables in total, most of which were unavailable. In our present study, data from 17 variables was extracted including marital status, sex, age at diagnosis, year of diagnosis, primary site, grade, ELN, tumor size, site-specific factor 1 (CEA), derived AJCC T and N, reason for no surgery, race recode, vital status record, SEER cause-specific death classification, survival month, total number of tumors for patients, and chemotherapy. The targeted population were patients diagnosed with stage II or III colon cancer between 2004 and 2010 according to the 6th edition of the AJCC cancer staging manual, received radical surgery, and survived more than 1 month. Exclusion criteria were as follows: patients with unavailable data of ELN, histology grade, and tumor size (Fig. 1). The number of ELN ranged from 0 to 90 and were divided into 3 groups of 0–10, 11–20, and more than 20 which can serve as a classified variable. Similarly, some continuous variables, including tumor size and age, were also regrouped as shown in the results.

Three cohorts were selected from datasets 1 and 2 for different purposes (Fig. 1). Cohort 1 from dataset 1 consisted of stage II colon cancer patients and was used to generate a novel prognostic scoring system that can be used for recognizing high-risk patients. Cohorts 2 and 3 were both selected from dataset 2. Cohort 2 consisted of stage II colon cancer patients while cohort 3 consisted of stage III colon cancer patients. Cohort 2 was used to validate the efficiency of the prognostic scoring system externally. A comparison between the novel prognostic scoring system in cohort 2 and TNM staging classification in cohort 3 was performed.

### Novel prognostic scoring system

For stage II colon cancer, the number of ELN was taken into consideration to regroup patients and tumor invasion depth. We defined the score as 0 or 1 when the tumor invasion depth was T3 or T4, respectively. Correspondingly, a score of 0, 1, or 2 indicated the number of ELN was more than 20, ranged from 11 to 20, or ranged from 0 to 10, respectively. All patients were reclassified into four groups based on the combined score, which was calculated as the sum of the scores (Table S1). Group N includes all the patients with a combined score N ( $N = 0, 1, 2, 3$ ).

### Statistical analyses

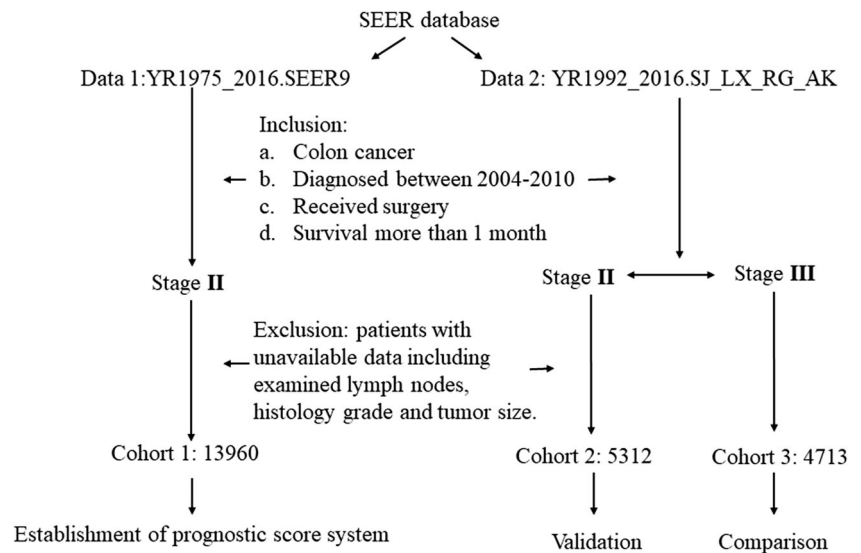
The difference between the groups was analyzed by  $\chi^2$  test. Univariate analysis was performed to evaluate the relationship between patients' survival and variables including marital status, sex, age, and tumor location. Multivariate Cox regression analysis was performed to identify the potential independent prognostic factors from the variables examined with  $P$  value  $< 0.05$  in univariate analysis. Survival curves were plotted by the Kaplan-Meier method. Hazard ratio and 95% confidence interval among subgroups were calculated by the log-rank test.  $\chi^2$  test, univariate, and multivariate analyses were performed using the Statistical Package for Social Sciences version 25. Survival difference between groups was compared using GraphPad Prism version 8. A two-sided  $P < 0.05$  was considered statistically significant.

## Results

### Patients' demographics and clinical characteristics

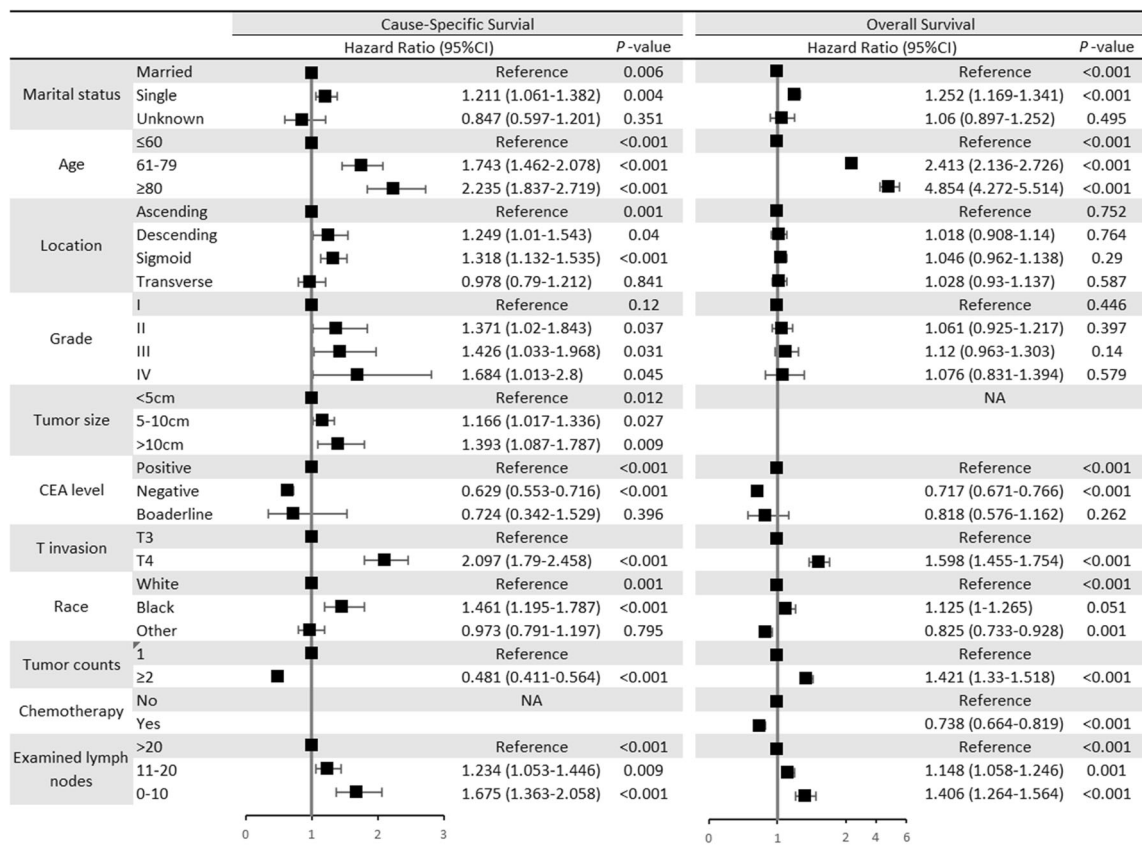
Three cohorts, delineated on a workflow chart (Fig. 1), were selected from the SEER database. After quality control analysis and filtering of data, patients were grouped into different cohorts. Cohort 1 from dataset 1 included a total of 13,960 patients with stage II colon cancer. Cohorts 2 and 3, both of

**Fig. 1** Flowchart for data selection



**Table 1** Demographics and clinical characteristics of cohort 1

Characteristics		0–10		11–20		> 20		P value
		N	%	N	%	N	%	
Marital status	Married	1056	49.50	4342	51.60	1814	53.20	0.096
	Single	974	45.60	3691	43.80	1435	42.10	
	Unknown	105	4.90	385	4.60	158	4.60	
Age	≤ 60	335	15.70	1671	19.90	922	27.10	< 0.001
	61–79	1054	49.40	3998	47.50	1618	47.50	
	≥ 80	746	34.90	2749	32.70	867	25.40	
Gender	Male	1030	48.20	3941	46.80	1604	47.10	0.498
	Female	1105	51.80	4477	53.20	1803	52.90	
Location	Ascending	957	44.80	4959	58.90	1668	49.00	< 0.001
	Descending	254	11.90	771	9.20	402	11.80	
	Sigmoid	650	30.40	1788	21.20	827	24.30	
	Transverse	274	12.80	900	10.70	510	15.00	
Grade	I	211	9.90	596	7.10	228	6.70	< 0.001
	II	1519	71.10	6160	73.20	2500	73.40	
	III	365	17.10	1498	17.80	595	17.50	
	IV	40	1.90	164	1.90	84	2.50	
Tumor size	< 5 cm	1372	64.30	4300	51.10	1437	42.20	< 0.001
	5–10 cm	705	33.00	3651	43.40	1749	51.30	
	> 10 cm	58	2.70	467	5.50	221	6.50	
CEA	Positive	383	36.10	1760	37.20	738	37.00	0.578
	Negative	666	62.80	2936	62.00	1246	62.50	
	Borderline	11	1.00	41	0.90	11	0.60	
T invasion	T3	1851	86.70	7416	88.10	3062	89.90	0.001
	T4	284	13.30	1002	11.90	345	10.10	
Tumor counts	1	1298	60.80	5410	64.30	2278	66.90	0.001
	≥ 2	837	39.20	3007	35.70	1129	33.10	
Chemotherapy	No	1730	81.00	7032	83.50	2812	82.50	0.018
	Yes	405	19.00	1386	16.50	595	17.50	
Race	White	1752	82.10	6848	81.60	2764	81.50	< 0.001
	Black	249	11.70	786	9.40	295	8.70	
	Other	133	6.20	761	9.10	334	9.80	



**Fig. 2** Multivariate survival analysis in cohort 1. Hazard ratio and 95% CI shown as forest plots. Those factors including married status, younger age at diagnosis, negative CEA, T3 stage, and increasing number of ELN were favorable prognostic factors for both CSS and OS in cohort 1

which were selected from dataset 2, included 5312 stage II and 4713 stage III colon cancer patients, respectively.

Demographics and clinical characteristics among the cohorts are listed in Table S2. All three cohorts had similar distributions of marital status, gender, age, and tumor location. Nearly half of the population analyzed consisted of patients diagnosed between ages 61 and 79. Patients with ascending colon cancer accounted for around fifty percent of patients across cohorts (Table S2). Cohorts 1 and 2, both of which only included stage II colon cancer patients, had similar distributions of histology grade, tumor size, CEA level, tumor counts, and chemotherapy (Table S2).

### Establishment of a novel prognostic scoring system

Cohort 1 was divided into 3 subgroups according to the number of examined regional lymph nodes. The differences in clinical characteristics among the three subgroups are listed in Table 1. There were no differences in marital status, gender, and CEA observed. Differences in features were small (less than 5%) but statistically significant ( $P < 0.05$ ) except for age, tumor location, and tumor size. Compared to group 0–10, group > 20 had a higher percentage of patients diagnosed at age  $\leq 60$  years old and a lower

percentage of patients diagnosed at age  $\geq 80$ . The proportion of patients with smaller tumor size (< 5 cm) was highest in group 0–10 and second highest in group 11–20 (Table 1).

Univariate and multivariate analyses were performed to evaluate the relationship between clinical characteristics and patients' survival. By univariate analysis, factors including marital status, age at diagnosis, tumor location, tumor histological grade, CEA level, T invasion, race, tumor counts, and ELN were found to be associated with both cause-specific survival (CSS) and overall survival (OS) of patients ( $P < 0.05$ ). Although tumor size was associated with CSS significantly ( $P < 0.001$ ), it had no association with OS ( $P = 0.102$ ). Interestingly, patients who received chemotherapy had a better OS but not CSS than those that did not receive chemotherapy ( $P < 0.001$ ) (Table S3). Factors with  $P < 0.05$  in univariate analysis were then further examined via multivariate analysis to identify independent prognostic factors. Results from the multivariate analysis are shown in Fig. 2. We found that independent prognostic factors for both CSS and OS included age at diagnosis, T invasion, number of ELN, and CEA level. Intriguingly, when compared to patients with only one primary tumor, those who had more than 2 tumors

showed better CSS ( $P < 0.001$ ; HR: 0.439, 95% CI: 0.390–0.495) but inferior OS ( $P < 0.001$ ; HR: 1.583, 95% CI: 1.509–1.660).

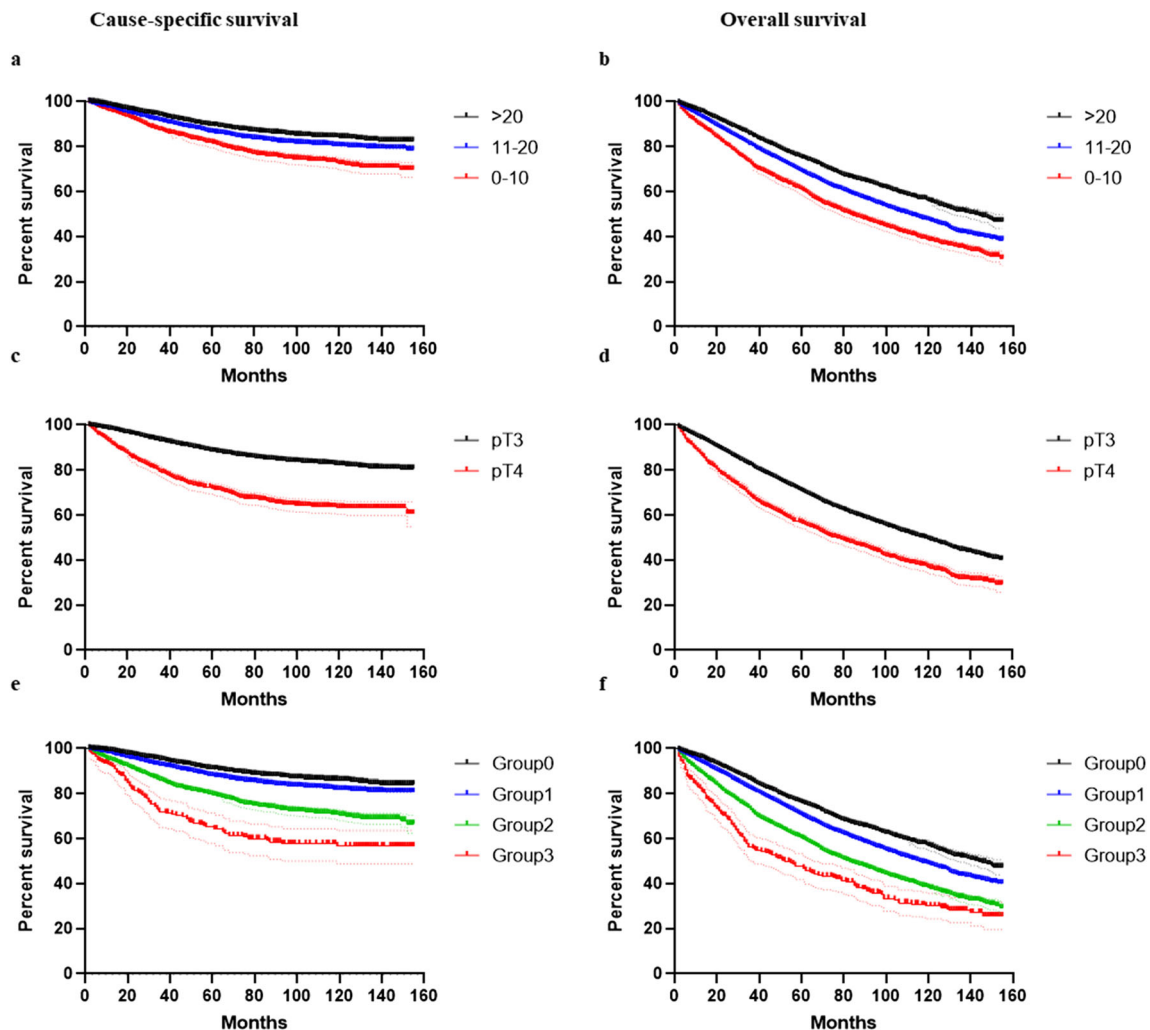
The survival curves for T invasion and ELN are illustrated in Fig. 3. Patients with a smaller number of ELN had both inferior CSS and OS (Fig. 3a, b). Patients with T4 tumors had both much worse CSS and OS when compared to T3 tumors (Fig. 3c, d). Furthermore, age at diagnosis was found to be an independent prognostic predictor and was distributed differently among the three groups. Survival curves were replotted after age was taken as a stratification factor. A smaller number of ELN predicted worse survival in subgroups of patients with different age ranges (Fig. S1).

A novel prognostic scoring system was established, as described in the methods section, by combining the T invasion and ELN to identify high-risk patients among stage II colon cancer patients who were considered as local tumors and had relatively better survival. As expected, patients with higher scores had worse survival (Fig. 3e, f).

## External validation of the novel prognostic scoring system

To validate the efficiency of the novel prognostic scoring system derived from cohort 1, cohort 2 was divided into four groups according to the new system. Univariate and multivariate analyses were then performed. The difference of distribution by clinical characteristics among the four groups is listed in Table 2. No difference was found for marital status and gender. Compared to the other groups, group 3 had a higher percentage of older patients, patients with positive CEA, or receiving chemotherapy ( $P < 0.001$ ). (More details can be found in Table 2.)

By univariate analysis, we found that T invasion, the number of ELN, and the combined score were associated with both CSS and OS ( $P < 0.001$ ). Patients who had more than one primary tumor showed better CSS and worse OS when compared to those with only one tumor, which coincided with the results from cohort 1. Additionally, chemotherapy only improved patients' OS but not CSS, which



**Fig. 3** Survival curves grouped by different factors in cohort 1. **a, c, e** Patients with smaller ELN numbers, pT4, and higher scores had worse cause-specific survival. **b, d, f** Patients with smaller ELN numbers, pT4, and higher scores had worse overall survival

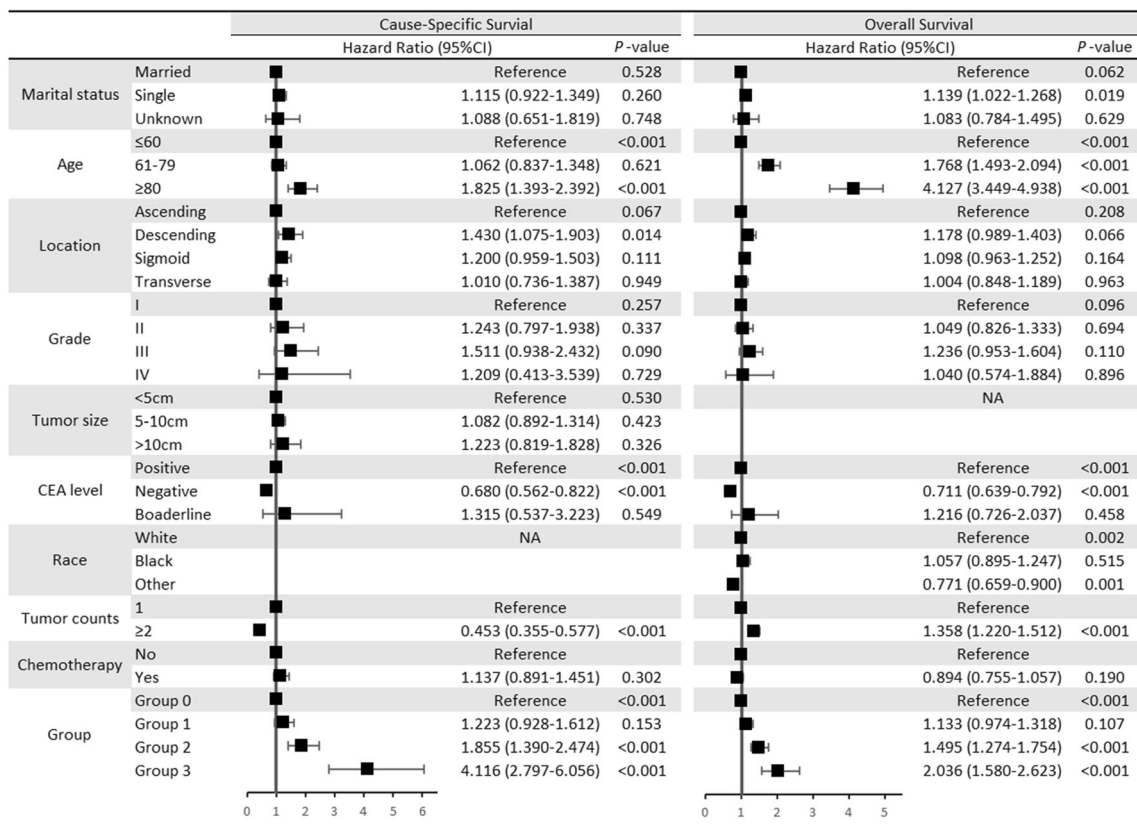
was also the case in cohort 1 (Table S4). Similarly, factors with  $P < 0.05$ , as determined by univariate analysis, except T invasion and ELN, which were used to generate the combination score, were included in the Cox multivariate model for further analysis. The combination score, patients' age, and CEA were all independent prognostic predictors in cohort 2 ( $P < 0.001$ ) (Fig. 4). Patients with higher scores had worse CSS and OS. Compared to group 0, the HR (95% CI) values of CSS, as estimated by log-rank test, were 1.223 (0.928–1.612), 1.855 (1.390–2.474), and 4.116 (2.797–6.056) for groups 1, 2, and 3, respectively. The HR (95% CI) values of OS were 1.495 (1.274–1.754) and 2.036 (1.580–2.623) for groups 2 and 3, respectively (Fig. 4). Thus, according to the novel prognostic scoring system, patients could be effectively classified into different risk levels.

### Clinical value of the novel prognostic scoring system

Stage III colon cancer is a regional advanced tumor accompanied by regional lymph node metastasis, which has a worse survival outcome than stage II colon cancer. Cohort 3 was disbanded into five subgroups according to T and N, which was determined following the 6th AJCC guidelines. Survival curves for subgroups are illustrated in Fig. 5a–d. The cumulative survival proportions at 3, 5, and 10 years after diagnosis are shown in Fig. 5e, f. Group 3 displayed the worst CSS and OS in cohort 2 (Fig. 5a, c), with comparable CSS and OS at 3, 5, and 10 years when compared with T3N2 or T4N1 groups (Fig. 5e, f; the hazard ratios for CSS (Table 3) and OS (Table 4) between the two groups were calculated and compared by the log-rank test.). Similar or worse survival groups, as determined by the novel scoring system and compared to

**Table 2** Demographics and clinical characteristics of cohort 2

Characteristics		Group 0		Group 1		Group 2		Group 3		P value
		N	%	N	%	N	%	N	%	
Marital status	Married	628	55.00	1289	52.50	782	51.90	110	52.40	0.583
	Single	478	41.90	1090	44.40	685	45.50	92	43.80	
	Unknown	35	3.10	76	3.10	39	2.60	8	3.80	
Age	≤ 60	365	32.00	607	24.70	289	19.20	39	18.60	< 0.001
	61–79	516	45.20	1141	46.50	751	49.90	91	43.30	
	≥ 80	260	22.80	707	28.80	466	30.90	80	38.10	
Gender	Male	533	46.70	1177	47.90	734	48.70	94	44.80	0.600
	Female	608	53.30	1278	52.10	772	51.30	116	55.20	
Location	Ascending	690	60.50	1338	54.50	638	42.40	91	43.30	< 0.001
	Descending	96	8.40	257	10.50	192	12.70	25	11.90	
	Sigmoid	230	20.20	611	24.90	471	31.30	71	33.80	
	Transverse	125	11.00	249	10.10	205	13.60	23	11.00	
Grade	I	51	4.50	158	6.40	98	6.50	18	8.60	0.033
	II	869	76.20	1882	76.70	1123	74.60	150	71.40	
	III	206	18.10	389	15.80	276	18.30	41	19.50	
	IV	15	1.30	26	1.10	9	0.60	1	0.50	
Tumor size	< 5 cm	457	40.10	1235	50.30	893	59.30	90	42.90	< 0.001
	5–10 cm	614	53.80	1104	45.00	566	37.60	94	44.80	
	> 10 cm	70	6.10	116	4.70	47	3.10	26	12.40	
CEA	Positive	222	32.80	498	34.00	322	38.30	63	51.20	0.002
	Negative	449	66.40	957	65.40	511	60.80	60	48.80	
	Borderline	5	0.70	9	0.60	7	0.80	0	0.00	
T invasion	T3	1141	100.00	2296	93.50	1191	79.10	0	0.00	< 0.001
	T4	0	0.00	159	6.50	315	20.90	210	100.00	
Tumor counts	1	785	68.80	1693	69.00	948	62.90	142	67.60	0.001
	≥ 2	356	31.20	762	31.00	558	37.10	68	32.40	
Chemotherapy	No	966	84.70	2112	86.00	1241	82.40	156	74.30	< 0.001
	Yes	175	15.30	343	14.00	265	17.60	54	25.70	
Race	White	849	74.70	1795	73.40	1050	69.80	138	65.70	0.001
	Black	95	8.40	270	11.00	194	12.90	24	11.40	
	Other	192	16.90	380	15.50	260	17.30	48	22.90	



**Fig. 4** Multivariate survival analysis in cohort 2. Hazard ratio and 95% CI shown as forest plots. Those factors including older age at diagnosis, positive CEA, higher score were inferior prognostic factors for both CSS and OS in cohort 2

TNM staging system scores, are italicized. Because right hemicolectomy usually yields a higher number of lymph nodes and the right colon cancer takes a majority of both cohorts. We performed a subgroup analysis, and similar results were also found in the subgroup of right colon cancer (Table S5 and S6). Based on our results, the novel prognostic scoring system could effectively identify high-risk patients with stage II colon cancer, who had an even worse survival than some of the stage III colon cancer. This, in turn, could be

applied in the clinical setting to help improve the prognosis of high-risk patients.

### Discussion

According to the AJCC cancer staging manual, primary tumor invasion (pT), involved regional lymph nodes (pN), and distant metastasis (pM) should be taken into consideration for

**Table 3** Hazard ratio calculated by comparison of CSS curves between groups by log-rank test

HR (95% CI)	Stage classification according to AJCC cancer staging manual, 6th edition				
	pT1-2N+ (Ref.)	pT3N1 (Ref.)	pT3N2 (Ref.)	pT4N1 (Ref.)	pT4N2 (Ref.)
Group 0	0.703 (0.511–0.967) *	0.333 (0.282–0.391) ***	0.201 (0.174–0.237) ***	0.170 (0.126–0.228) ***	0.118 (0.084–0.165) ***
Group 1	0.998 (0.769–1.294) <sup>N</sup>	0.470 (0.409–0.539) ***	0.282 (0.241–0.329) ***	0.235 (0.176–0.313) ***	0.158 (0.111–0.221) ***
Group 2	1.541 (1.215–1.955) **	0.731 (0.632–0.844) ***	0.440 (0.379–0.510) ***	0.369 (0.289–0.471) ***	0.250 (0.188–0.330) ***
Group 3	3.842 (2.619–5.636) ***	1.851 (1.368–2.505) ***	1.103 (0.861–1.412) <sup>N</sup>	0.932 (0.714–1.216) <sup>N</sup>	0.629 (0.487–0.812) **

<sup>N</sup>No statistical significance

\*  $P < 0.05$

\*\*  $P < 0.01$

\*\*\*  $P < 0.001$

**Table 4** Hazard ratio calculated by comparison of OS curves between groups by log-rank test

HR (95% CI)	Stage classification according to AJCC cancer staging manual, 6 <sup>th</sup> edition				
	pT1-2N+ (Ref.)	pT3N1 (Ref.)	pT3N2 (Ref.)	pT4N1 (Ref.)	pT4N2 (Ref.)
Group 0	0.802 (0.682–0.944) **	0.605 (0.545–0.670) ***	0.459 (0.409–0.515) ***	0.391 (0.323–0.474) ***	0.278 (0.224–0.343) ***
Group 1	1.019 (0.888–1.168) <sup>N</sup>	0.766 (0.704–0.833) ***	0.576 (0.521–0.637) ***	0.490 (0.410–0.584) ***	0.337 (0.274–0.413) ***
Group 2	1.405 (1.234–1.600) ***	1.065 (0.974–1.165) <sup>N</sup>	0.805 (0.729–0.888) ***	0.692 (0.592–0.809) ***	0.484 (0.405–0.577) ***
Group 3	2.122 (1.678–2.685) ***	1.624 (1.322–1.996) ***	1.211 (1.005–1.459) *	1.055 (0.861–1.292) <sup>N</sup>	0.754 (0.620–0.916) **

<sup>N</sup>No statistical significance

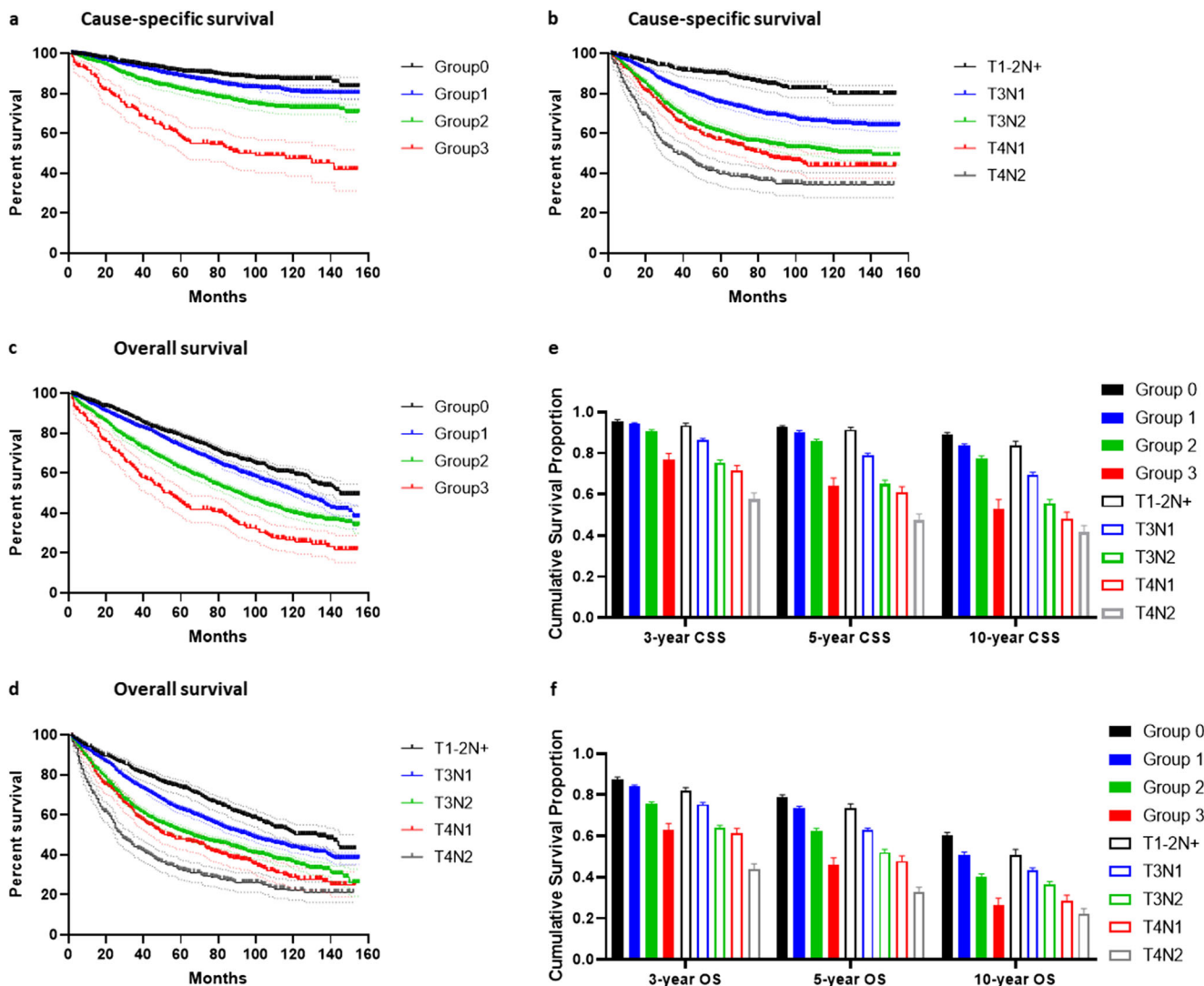
\*  $P < 0.05$

\*\*  $P < 0.01$

\*\*\*  $P < 0.001$

colon cancer staging, and stage II colon cancer includes only T3–4 tumor without regional lymph node consideration.

There is a big difference in treatment between stage II and III colon cancer. Precise assessment of stage is very important



**Fig. 5** Comparison of survival curves between reclassified stage II patients using the novel system and stage III patients. **a, c** Survival outcome of patients regrouped using the novel prognostic scoring

system in cohort 2. **b, d** Survival outcome of subgroups in stage III patients of cohort 3. **e, f** Comparison of long-term survival of CSS and OS among subgroups of cohorts 2 and 3



for the delivery of adjuvant therapy to colon cancer patients. The pN stage in colon cancer is influenced by regional lymph nodes. Adequate number of ELN for pN staging is no less than 12 as recommended by AJCC guidelines. Inadequate ELN numbers may result in tumor understaging and lead to worse survival [12, 13, 16–18]. However, the optimal number of ELN has not been determined in colon cancer [19–24]. Here, we focused on the impact of ELN number on stage II colon cancer and investigated the stage migration effect caused by insufficient ELN number.

Several factors influence the number of ELN. Operation selection and surgeon skill influence complete regional lymph node resection, while variations in pathology practice and skill lead to different regional nodes retrieved [20, 25, 26]. In our present study, some patient characteristics were associated with the number of ELN. Younger patients or those with bigger tumors were more likely to have higher lymph node yield. Previous studies have evaluated the impact of ELN number and metastatic lymph node ratio (mLNR, calculated by positive lymph nodes dividing ELN number) on the prognosis of cancer patients and confirmed higher ELN associated with better survival across various solid tumors including colon cancer [13, 20, 22, 27–29]. In our study, ELN number, pT, and patients' age were determined to be independent prognostic predictors in stage II colon cancer. Patients with less than 10 ELN exhibited worse CSS and OS as compared to those with higher ELN. Similar results were observed among subgroups with different age ranges.

Since the method for building a nomogram for cancer prognosis was introduced by Alexia Iasonos et al. in 2008 [30], studies on the establishment of prognostic nomograms for different cancers have expanded [31–35]. Although useful for stratifying patients, nomograms, especially those consisting of many factors, have limitations for clinical practice. Our novel prognostic scoring system, consisting of combining pT and ELN, is easy and convenient for clinical practice. Stage II colon cancer patients can be reclassified by simply adding the ELN number. As discussed above, patients with the highest score had the worst survival of CSS and OS in both training and validation cohorts according to the novel scoring system.

Although patients with stage II colon cancer generally have a better prognosis than those with stage III colon cancer, there is still extensive variability. One possible reason for this is that stage II colon cancer patients might be stage III colon cancer patients, suggesting understaging as a result of inadequate ELN. The impact of the number of ELN on accurate staging in different cancers has been assessed and generally, a greater number of ELN is related to more accurate node staging and better prognosis [9, 10, 36]. To evaluate the stage migration effect caused by inadequate ELN in stage II colon cancer, we compared patient survival of reclassified stage II colon cancer with those of stage III colon cancer. As expected, stage II

colon cancer patients with a smaller number of ELN had comparable, or worse, survival of both CSS and OS when compared with stage III colon cancer patients. For instance, pT4N0 cancer with a 0–10 ELN had similar survival of both CSS and OS relative to either pT4N1 or pT3N2 cancer. Inadequate ELN leads to a greater chance of failing to detect positive lymph nodes and to understaging in colon cancer.

Our present study focused on the effect of inadequate ELN on the prognosis of patients without regional lymph node metastasis. However, the limitation is that why patients experienced inadequate ELN was not addressed here. Some molecular markers such as microsatellite instability (MSI), KRAS, and BRAF mutations are not documented which can also predict the prognosis and direct the treatment of specific stage CRC. MSI test is universally recommended for diagnosis of Lynch syndrome and also can direct the adjuvant chemotherapy post-surgery for a stage II population (irrespective of germline or sporadic background) [37, 38]. However, identification of gene mutations like KRAS and BRAF is recommended for metastatic colon cancer in clinical practice.

In conclusion, a greater number of ELN were found to be associated with better survival in stage II colon cancer and we recommend more than 20 ELN for accurate staging post-surgery. According to the prognostic scoring system, stage II colon cancer can be reclassified into different subgroups, where patients with a score of 2 or 3 should be treated following guidelines for stage III colon cancer patients.

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1007/s00384-020-03794-6>.

**Acknowledgments** We thank Darleny Lizardo (Department of Pharmacology and Chemical Biology, UPMC Hillman Cancer Centre, Pittsburgh) for critically reading the manuscript. The authors acknowledge the efforts of the Surveillance, Epidemiology, and End Results (SEER) Program in the creation and maintenance of the SEER database (<https://seer.cancer.gov/>).

**Author contributions** Xiangping Song and Di Xie carried out the concept and designed the study. Xiangping Song, Di Xie, and Lingling Tong had the responsibility for the integrity and accuracy of data analysis. Di Xie and Xiangping Song wrote the manuscript and all the authors approved the final version.

**Funding** This work was supported by the China Scholarship Council (201806370121).

**Data availability** The data sets used and/or analyzed during the current study are available from the SEER database.

## Compliance with ethical standards

**Conflict of interest** The authors declare that they have no conflict of interest.

**Ethical statement** SEER is a publicly available database with anonymized data; no ethical review was required.

## References

- Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A (2018) Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin* 68(6):394–424. <https://doi.org/10.3322/caac.21492>
- Siegel RL, Miller KD, Fedewa SA, Ahnen DJ, Meester RGS, Barzi A, Jemal A (2017) Colorectal cancer statistics, 2017. *CA Cancer J Clin* 67(3):177–193. <https://doi.org/10.3322/caac.21395>
- Schmoll HJ, Van Cutsem E, Stein A, Valentini V, Glimelius B, Haustermans K, Nordlinger B, van de Velde CJ, Balmana J, Regula J, Nagtegaal ID, Beets-Tan RG, Arnold D, Ciardiello F, Hoff P, Kerr D, Köhne CH, Labianca R, Price T, Scheithauer W, Sobrero A, Tabernero J, Aderka D, Barroso S, Bodoky G, Douillard JY, El Ghazaly H, Gallardo J, Garin A, Glynne-Jones R, Jordan K, Meshcheryakov A, Papamichail D, Pfeiffer P, Souglakos I, Turhal S, Cervantes A (2012) ESMO Consensus Guidelines for management of patients with colon and rectal cancer. a personalized approach to clinical decision making. *Ann Oncol* 23(10):2479–2516. <https://doi.org/10.1093/annonc/mds236>
- Hashiguchi Y, Muro K, Saito Y, Ito Y, Ajioka Y, Hamaguchi T, Hasegawa K, Hotta K, Ishida H, Ishiguro M, Ishihara S, Kanemitsu Y, Kinugasa Y, Murofushi K, Nakajima TE, Oka S, Tanaka T, Taniguchi H, Tsuji A, Uehara K, Ueno H, Yamanaka T, Yamazaki K, Yoshida M, Yoshino T, Itabashi M, Sakamaki K, Sano K, Shimada Y, Tanaka S, Uetake H, Yamaguchi S, Yamaguchi N, Kobayashi H, Matsuda K, Kotake K, Sugihara K (2020) Japanese Society for Cancer of the Colon and Rectum (JSCCR) guidelines 2019 for the treatment of colorectal cancer. *Int J Clin Oncol* 25(1):1–42. <https://doi.org/10.1007/s10147-019-01485-z>
- O'Connor ES, Greenblatt DY, LoConte NK, Gangnon RE, Liou JJ, Heise CP, Smith MA (2011) Adjuvant chemotherapy for stage II colon cancer with poor prognostic features. *J Clin Oncol* 29(25):3381–3388. <https://doi.org/10.1200/jco.2010.34.3426>
- Gkekak I, Novotny J, Fabian P, Nemecek R, Palmqvist R, Strigård K, Pecen L, Svoboda T, Gurlich R, Gunnarsson U (2019) Deficient mismatch repair as a prognostic marker in stage II colon cancer patients. *Eur J Surg Oncol* 45(10):1854–1861. <https://doi.org/10.1016/j.ejso.2019.05.023>
- Gkekak I, Novotny J, Fabian P, Nemecek R, Palmqvist R, Strigård K, John S, Pecen L, Reginacova K, Gunnarsson U (2019) Mismatch repair status predicts survival after adjuvant treatment in stage II colon cancer patients. *J Surg Oncol* 121:392–401. <https://doi.org/10.1002/jso.25798>
- Auclin E, André T, Taieb J, Banzi M, Van Laethem JL, Tabernero J, Hickish T, de Gramont A, Vernerey D (2019) Association of post-operative CEA with survival and oxaliplatin benefit in patients with stage II colon cancer: a post hoc analysis of the MOSAIC trial. *Br J Cancer* 121(4):312–317. <https://doi.org/10.1038/s41416-019-0521-7>
- Huang L, Jansen L, Balavarca Y, van der Geest L, Lemmens V, Groot Koerkamp B, van Santvoort HC, Grützmann R, Besselink MG, Schrotz-King P, Brenner H (2019) Significance of examined lymph node number in accurate staging and long-term survival in resected stage I-II pancreatic cancer—more is better? A large international population-based cohort study. *Ann Surg*. <https://doi.org/10.1097/sla.0000000000003558>
- Liang W, He J, Shen Y, Shen J, He Q, Zhang J, Jiang G, Wang Q, Liu L, Gao S, Liu D, Wang Z, Zhu Z, Ng CS, Liu CC, Horsleben Petersen R, Rocco G, D'Amico T, Brunelli A, Chen H, Zhi X, Liu B, Yang Y, Chen W, Zhou Q, He J (2017) Impact of examined lymph node count on precise staging and long-term survival of resected non-small-cell lung cancer: a population study of the US SEER Database and a Chinese Multi-Institutional Registry. *J Clin Oncol* 35(11):1162–1170. <https://doi.org/10.1200/jco.2016.67.5140>
- Zhu X, Zhao M, Zhou L, Zhang M, Cao P, Tao L (2020) Significance of examined lymph nodes number and metastatic lymph nodes ratio in overall survival and adjuvant treatment decision in resected laryngeal carcinoma. *Cancer Med* 9:3006–3014. <https://doi.org/10.1002/cam4.2902>
- Tsai HL, Huang CW, Chen CW, Yeh YS, Ma CJ, Wang JY (2016) Survival in resected stage II colorectal cancer is dependent on tumor depth, vascular invasion, postoperative CEA level, and the number of examined lymph nodes. *World J Surg* 40(4):1002–1009. <https://doi.org/10.1007/s00268-015-3331-y>
- Trepanier M, Erkan A, Kouyoumdjian A, Nassif G, Albert M, Monson J, Lee L (2019) Examining the relationship between lymph node harvest and survival in patients undergoing colectomy for colon adenocarcinoma. *Surgery* 166(4):639–647. <https://doi.org/10.1016/j.surg.2019.03.027>
- Ahmed S, Leis A, Chandra-Kanthan S, Fields A, Zaidi A, Abbas T, Le D, Reeder B, Pahwa P (2016) Regional lymph nodes status and ratio of metastatic to examined lymph nodes correlate with survival in stage IV colorectal cancer. *Ann Surg Oncol* 23(7):2287–2294. <https://doi.org/10.1245/s10434-016-5200-9>
- Pyo JS, Shin YM, Kang DW (2019) Prognostic implication of metastatic lymph node ratio in colorectal cancers: comparison depending on tumor location. *J Clin Med* 8(11). <https://doi.org/10.3390/jcm8111812>
- Wang Y, Zhou M, Yang J, Sun X, Zou W, Zhang Z, Zhang J, Shen L, Yang L, Zhang Z (2019) Increased lymph node yield indicates improved survival in locally advanced rectal cancer treated with neoadjuvant chemoradiotherapy. *Cancer Med* 8(10):4615–4625. <https://doi.org/10.1002/cam4.2372>
- Hong T, Cai D, Jin L, Zhang Y, Lu T, Hua D, Wu X (2020) Development and validation of a nomogram to predict survival after curative resection of nonmetastatic colorectal cancer. *Cancer Med* 9:4126–4136. <https://doi.org/10.1002/cam4.3010>
- Chang GJ, Rodriguez-Bigas MA, Skibber JM, Moyer VA (2007) Lymph node evaluation and survival after curative resection of colon cancer: systematic review. *J Natl Cancer Inst* 99(6):433–441. <https://doi.org/10.1093/jnci/djk092>
- Cai Y, Cheng G, Lu X, Ju H, Zhu X (2020) The re-evaluation of optimal lymph node yield in stage II right-sided colon cancer: is a minimum of 12 lymph nodes adequate? *Int J Color Dis* 35(4):623–631. <https://doi.org/10.1007/s00384-019-03483-z>
- Chen SL, Bilchik AJ (2006) More extensive nodal dissection improves survival for stages I to III of colon cancer: a population-based study. *Ann Surg* 244(4):602–610. <https://doi.org/10.1097/01.sla.0000237655.11717.50>
- Peeples C, Shellnut J, Wasvary H, Riggs T, Sacksner J (2010) Predictive factors affecting survival in stage II colorectal cancer: is lymph node harvesting relevant? *Dis Colon Rectum* 53(11):1517–1523. <https://doi.org/10.1007/DCR.0b013e3181f20116>
- Vather R, Sasmour T, Kahokehr A, Connolly AB, Hill AG (2009) Lymph node evaluation and long-term survival in stage II and stage III colon cancer: a national study. *Ann Surg Oncol* 16(3):585–593. <https://doi.org/10.1245/s10434-008-0265-8>
- Del Paggio JC, Peng Y, Wei X, Nanji S, MacDonald PH, Krishnan Nair C, Booth CM (2017) Population-based study to re-evaluate optimal lymph node yield in colonic cancer. *Br J Surg* 104(8):1087–1096. <https://doi.org/10.1002/bjs.10540>
- Tsai HL, Lu CY, Hsieh JS, Wu DC, Jan CM, Chai CY, Chu KS, Chan HM, Wang JY (2007) The prognostic significance of total lymph node harvest in patients with T2–4N0M0 colorectal cancer. *J Gastrointest Surg* 11(5):660–665. <https://doi.org/10.1007/s11605-007-0119-x>

25. Lemmens VE, van Lijnschoten I, Janssen-Heijnen ML, Rutten HJ, Verheij CD, Coebergh JW (2006) Pathology practice patterns affect lymph node evaluation and outcome of colon cancer: a population-based study. *Ann Oncol* 17(12):1803–1809. <https://doi.org/10.1093/annonc/mdl312>
26. Reinbach DH, McGregor JR, Murray GD, O'Dwyer PJ (1994) Effect of the surgeon's specialty interest on the type of resection performed for colorectal cancer. *Dis Colon Rectum* 37(10):1020–1023. <https://doi.org/10.1007/bf02049316>
27. Zheng YZ, Li XQ, Wang JY, Yang H, Wen J, Zhai WY, Yuan LX, Fu SS, Liao HY, Fu JH (2020) Impact of examined lymph node count for oesophageal squamous cell carcinoma in patients who underwent left transthoracic oesophagectomy. *Eur J Surg Oncol* 46:1956–1962. <https://doi.org/10.1016/j.ejso.2020.04.047>
28. Martinez-Meehan D, Lutfi W, Dhupar R, Christie N, Baker N, Schuchert M, Luketich JD, Okusanya OT (2020) Factors associated with survival in complete pathologic response non-small cell lung cancer. *Clin Lung Cancer* 21:349–356. <https://doi.org/10.1016/j.clcc.2020.03.003>
29. Kano K, Yamada T, Yamamoto K, Komori K, Watanabe H, Hara K, Shimoda Y, Maezawa Y, Fujikawa H, Aoyama T, Tamagawa H, Yamamoto N, Cho H, Shiozawa M, Yukawa N, Yoshikawa T, Morinaga S, Rino Y, Masuda M, Ogata T, Oshima T (2020) Association between lymph node ratio and survival in patients with pathological stage II/III gastric cancer. *Ann Surg Oncol* 27:4235–4247. <https://doi.org/10.1245/s10434-020-08616-1>
30. Iasonos A, Schrag D, Raj GV, Panageas KS (2008) How to build and interpret a nomogram for cancer prognosis. *J Clin Oncol* 26(8):1364–1370. <https://doi.org/10.1200/jco.2007.12.9791>
31. Li Y, Zhang G, Song X, Zhao L, Güngör C, Wang D, Liu W, Huang Y, Tan F (2020) Establishment and verification of synchronous metastatic nomogram for gastrointestinal stromal tumors (GISTs): a population-based analysis. *Gastroenterol Res Pract* 2020:8493707–8493707. <https://doi.org/10.1155/2020/8493707>
32. Li C, Pei Q, Zhu H, Tan F, Zhou Z, Zhou Y, Li Y, Pei H (2018) Survival nomograms for stage III colorectal cancer. *Medicine* 97(49):e13239. <https://doi.org/10.1097/md.00000000000013239>
33. Zheng P, Lai C, Yang W, Guo J, Xiao S, Chen Z (2020) Nomogram predicting cancer-specific survival in elderly patients with stages I–III colon cancer. *Scand J Gastroenterol* 55(2):202–208. <https://doi.org/10.1080/00365521.2020.1720280>
34. Hu J, Li H, He T, Deng H, Gong G, Cui Y, Liu P, Ren W, Li C, Chen J, Zu X (2020) A nomogram incorporating PD-L1, NLR, and clinicopathologic features to predict inguinal lymph node metastasis in penile squamous cell carcinoma. *Urol Oncol* 38:641.e19–641.e29. <https://doi.org/10.1016/j.urolonc.2020.04.015>
35. Chen L, Zeng F, Yao L, Fang T, Liao M, Long J, Xiao L, Deng G (2020) Nomogram based on inflammatory indices for differentiating intrahepatic cholangiocarcinoma from hepatocellular carcinoma. *Cancer Med* 9(4):1451–1461. <https://doi.org/10.1002/cam4.2823>
36. Hua J, Zhang B, Xu J, Liu J, Ni Q, He J, Zheng L, Yu X, Shi S (2019) Determining the optimal number of examined lymph nodes for accurate staging of pancreatic cancer: an analysis using the nodal staging score model. *Eur J Surg Oncol* 45(6):1069–1076. <https://doi.org/10.1016/j.ejso.2019.01.018>
37. Dienstmann R, Villacampa G, Sveen A, Mason MJ, Niedzwiecki D, Nesbakken A, Moreno V, Warren RS, Lothe RA, Guinney J (2019) Relative contribution of clinicopathological variables, genomic markers, transcriptomic subtyping and microenvironment features for outcome prediction in stage II/III colorectal cancer. *Ann Oncol* 30(10):1622–1629. <https://doi.org/10.1093/annonc/mdz287>
38. Zhang CM, Lv JF, Gong L, Yu LY, Chen XP, Zhou HH, Fan L (2016) Role of deficient mismatch repair in the personalized management of colorectal cancer. *Int J Environ Res Public Health* 13(9). <https://doi.org/10.3390/ijerph13090892>

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.